

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 1999, 15:16:41 ; Search time 10.22 Seconds

(without alignments)  
926.603 Million cell updates/sec

Title: US-09-126-945-2

Perfect score: 1785

Sequence: 1 MGSA5PGLSVSPSHLLLP.....GIRKPDISORLYQVHP1 335

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	354	19.8	114	1	ETS4_DROME	P29775 drosophila
2	300.5	16.8	462	1	ERG_HUMAN	P11308 homo sapien
3	294	16.5	464	1	ERG_DROME	Q06688 drosophila
4	292	16.4	478	1	ERG_CHICK	Q08837 gallus gall
5	289	16.2	454	1	GABA_HUMAN	Q06346 homo sapien
6	288	16.1	669	1	MYBE_AVI	P01105 avian leuk
7	277	15.5	454	1	GABA_MOUSE	Q00422 mus musculu
8	275	15.4	453	1	FLI1_XENLA	P41157 xenopus lae
9	272	15.2	485	1	ETS2_CHICK	P15063 gallus gall
10	272	15.2	452	1	FLI1_MOUSE	P26323 mus musculu
11	270	15.1	452	1	FLI1_HUMAN	Q01343 homo sapien
12	265.5	14.9	485	1	ETV6_MOUSE	P97360 mus musculu
13	259.5	14.5	452	1	ETV6_HUMAN	P41212 homo sapien
14	252.5	14.1	441	1	ETS2_CHICK	P13474 gallus gall
15	250	14.0	472	1	ETV2_XENLA	Q91712 xenopus lae
16	247.5	13.9	829	1	ETV4_DROME	P20105 drosophila
17	247.5	13.9	883	1	ETV4_DROME	P11336 drosophila
18	246.5	13.8	469	1	ETS2_HUMAN	P15036 homo sapien
19	246	13.8	468	1	ETS2_MOUSE	P15037 mus musculu
20	244.5	13.7	438	1	ETV1_XENLA	P18755 xenopus lae
21	243.5	13.6	440	1	ETS1_MOUSE	P27577 mus musculu
22	243	13.6	472	1	ETS2_XENLA	P19102 xenopus lae
23	241.5	13.5	441	1	ETS1_HUMAN	P14921 homo sapien
24	240.5	13.5	441	1	ETS1_RAT	P11556 ratius norv
25	237	13.3	479	1	ETS2_CHICK	P10157 gallus gall
26	236	13.2	510	1	ERG_HUMAN	P41161 homo sapien
27	233.5	13.1	619	1	ELF1_HUMAN	P32519 homo sapien
28	233.5	13.1	612	1	ELF1_MOUSE	Q06775 mus musculu
29	224	12.5	428	1	ELK1_HUMAN	P19419 homo sapien
30	224	12.5	429	1	ELK1_MOUSE	P19466 mus musculu
31	224	12.5	555	1	ETV4_MOUSE	P28332 mus musculu
32	222.5	12.5	732	1	POK_DROME	Q01842 drosophila
33	222	12.4	551	1	ETV4_HUMAN	P43268 homo sapien
34	217.5	12.2	477	1	ETV1_HUMAN	P28545 homo sapien
35	216.5	12.1	431	1	SAPB_HUMAN	P28323 homo sapien
36	216.5	12.1	405	1	SAPB_MOUSE	P18164 mus musculu
37	212.5	11.9	477	1	ETV1_MOUSE	P41158 mus musculu
38	210.5	11.8	430	1	SAPB_MOUSE	P41158 mus musculu
39	208.5	11.7	431	1	ETV2_HUMAN	Q00321 homo sapien
40	205.5	11.5	623	1	PNT1_DROME	P51022 drosophila
41	205.5	11.5	718	1	PNT2_DROME	P51023 drosophila
42	200.5	11.2	407	1	ELK3_HUMAN	P41970 homo sapien
43	199	11.1	250	1	ETV3_HUMAN	P41162 homo sapien

44 198.5 11.1 409 1 ELK3\_MOUSE  
45 197.5 11.1 268 1 ETV1\_XENLA

P41971 mus musculu  
P18756 xenopus lae

## ALIGNMENTS

```

RESULT 1
ETS4_DROME STANDARD; PRT; 114 AA.
AC P29775;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE DNA-BINDING PROTEIN D-ETS-4 (FRAGMENT).
GN ETS98B OR ETS-4.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RM SEQUENCE FROM N.A.
RA MEDLINE; 92249640.
RA CHEN T., BURNING M., KARIM F.D., THUMMEL C.S.;
RT "Isolation and characterization of five drosophila genes that encode
an ets-related DNA binding domain."
RL DEV. BIOL. 151:176-191(1992).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M88474; G157196; -
CC PIR: S28821; S28821; Ets98B.
CC FLYBASE: FBgn005659; Ets98B.
CC DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
CC DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
CC DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
CC PRFAM: PF00178; Ets; 1.
CC DR HSP: Q01543; 1FLI.
CC KW DNA-BINDING; NUCLEAR PROTEIN.
CC FT NON_TER 1 1
CC FT DNA_BIND 26 108 ETS-DOMAIN.
CC FT NON_TER 114 114
CC SQ SEQUENCE 114 AA; 13051 MW; E9003913 CRC32;

Query Match 19.8%; Score 354; DB 1; Length 114;
Best Local Similarity 74.2%; Pred. No. 4.3e-21;
Matches 66; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

OY 246 GPHLHMOFLKELLKHSYGRFIRMLNKEGIRKEDSAOYALNKRKRPAMNTDKL 305
DB 23 GSHLHMOFLKELLASPDVNGTAIRWIDRSKIGIEDSVAVAKLWGRKRPAMNTDKL 82
OY 306 SRSIRORYKGIKIRKPDISORLYQVHP 334
DB 83 SRSIRORYKGIKIRKPDISORLYQVHP 111

RESULT 2
ERG_HUMAN STANDARD; PRT; 462 AA.
AC P11308;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

```

DE 1) TRANSCRIPTIONAL REGULATOR ERG (TRANSFORMING PROTEIN ERG) (ERG-2/ERG-  
GN ENG.  
OC HOMO SAPIENS (HUMAN).  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA.  
OC PRIMATES: CATARRHINI: HOMINIDAE: HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 87263429.  
RA RAO V.N., PAPAS T.S., SHYAM E., REDDY P.;  
RT "erg, a human ets-related gene on chromosome 21: alternative  
RT splicing, polyadenylation, and translation.";  
RL SCIENCE 237:635-639(1987).  
RN [2]  
RP SEQUENCE OF 100-462 FROM N.A.  
RX MEDLINE: 87317608.  
RA REDDY E.S.P., RAO V.N., PAPAS T.S.;  
RT "The erg gene: a human gene related to the ets oncogene.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 84:6131-6135(1987).  
RN [3]  
RP CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE: 94356859.  
RA DUNN T., PRAISSMAN L., HAGAG N., VIOLA M.V.;  
RT "ERG gene is translocated in an Ewing's sarcoma cell line.";  
RL CANCER GENET. CYTOGENET. 76:19-22(1994).  
RN [4]  
RP CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE: 94243799.  
RA ICHIKAWA H., SHIMITSU K., HAYASHI Y., OHKI M.;  
RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human  
RL myeloid leukemia with t(16;21) chromosomal translocation.";  
RX CANCER RES. 54:2865-2868(1994).  
CC -1- FUNCTION: ACTS AS A TRANSCRIPTIONAL ACTIVATOR (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED  
CC BY A CHROMOSOMAL TRANSLOCATION T(16;21)(P11;Q22) WHICH INVOLVES  
CC ERG AND FUS.  
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M1754; G182187; -;  
DR EMBL: M21535; G182185; -;  
DR PIR: A29515; TVHUEG.  
DR MIM: 165080; -;  
DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
DR PIR: P00178; ETS; 1.  
DR HSP: Q01543; EFLI.  
DR TRANSFAC: T00265; -;  
DR TRANSFAC: T00266; -;  
KM TRANSFORMING REGULATION: ACTIVATOR, NUCLEAR PROTEIN, DNA-BINDING;  
KM CHROMOSOMAL TRANSLOCATION.  
FT CHAIN 1 462 TRANSFORMING PROTEIN ERG-2.  
FT DNA\_BIND 100 462 TRANSFORMING PROTEIN ERG-1.  
FT DNA\_BIND 294 374 ETS-DOMAIN.  
SQ SEQUENCE 462 AA; 52031 MW; CEC2D2B CRC32;

DR 6 PDPAHIKEALVSSEDSLFECAVGPPLAKTEMTASSSSDYGQTSKSPRYQDML 65  
QY 85 EEP-----EEPOCPVIDSOAPAGSLDLP--GGTLLEHSLDYOQ 123  
DB 66 QPARVATIKMECNSQVNSGRNSPDECSVAKGKMGSPDTVGMNYSWEKHM--PP 123  
QY 124 SMVGEVLKDEITACKLLNTADPMSPNSVOKMLNTEHOVRLPMG-KAPQELAGE 182  
DB 124 NMTNE-----RVIVPADPILMTSHVROMLEMAVKEGLDPVILLFONIDGE 174  
QY 183 LCMSEDFQRSD-LGDVLAHL-----DIWKS-----212  
DB 175 LCKMTKDFQRLPSPYNADILLSHLHRLREPLPLHLSIDVDKALQNSPRLHARTDLP 234  
QY 212 -----NAW-----NKERTSPGAIHYCASTEE-----SWTSEVDSGS 245  
DB 235 YEPFRSAGMTGHHPTPOSAAPSPSTV--PRTDQRPQDLPYQILQPTSSRLANPS 291  
QY 246 GQPIHMOFLKEDLLRPHSYGRFIRMLNKEGIFKIEDSAQVRLGIRKNPAMNDKL 305  
DB 292 GO-IOLWQFLLE-LTSDSSNSCITW-EGTNGEFKMTDPEVARNRGKSPNNYDKL 348  
QY 306 SRGIRYKKGITRKDISRLYQF 331  
DB 349 SRALRIYDKNMTKVH-GKRYAYKF 373

RESULT 3  
ELG\_DROME  
ID ELG\_DROME STANDARD: PRT: 464 AA.  
AC 004688;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE DNA-BINDING PROTEIN D-ELG.  
GN ETS97D OR ELG.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA: METAZOA: ARTHROPODA: TRACHEATA: HEXAPODA: INSECTA;  
OC PTERYGOTA: DIPTERA: BRACHYCERA: MOSCOMORPHA: EPHYDROIDEA;  
OC DROSOPHILINAE: DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93096481.  
RA THE S.M., XIE X., SMYTH F., PAPAS T.S., WATSON D.K., SCHULTZ R.A.;  
RT "Molecular characterization and structural organization of D-elig, an  
RT ets proto-oncogene-related gene of Drosophila.";  
RL ONCOGENE 7:2471-2478(1992).  
RN [2]  
RP SEQUENCE OF 323-463 FROM N.A.  
RX MEDLINE: 91319397.  
RA PRIBYL L.J., WATSON D.K., SCHULTZ R.A., PAPAS T.S.;  
RT "D-elig, a member of the Drosophila ets gene family: sequence,  
RT expression and evolutionary comparison.";  
RL ONCOGENE 6:1175-1183(1991).  
RN [3]  
RP SEQUENCE OF 298-449 FROM N.A.  
RX MEDLINE: 92249640.  
RA CHEN T., BUNTING M., KARIM F.D., THUMMEL C.S.;  
RT "Isolation and characterization of five Drosophila genes that encode  
RT an ets-related DNA binding domain.";  
RL DEV. BIOL. 151:176-191(1992).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X68259; G7943; -;

QY 45 PSPAPTEQGLSFR-----YSTYDML-----YPEDSSMAAKPGASRSRE 84  
Matches 103; Conservative 48; Mismatches 118; Indels 117; Gaps 17;



RP SEQUENCE FROM N.A.  
 RX MEDLINE: 95097980.  
 RA GUGNEJA S., VIRBASTIUS J.V., SCARPUOLA R.C.;  
 RT "Four structurally distinct, non-DNA-binding subunits of human  
 RT nuclear respiratory factor 2 share a conserved transcriptional  
 RT activation domain."  
 RL MOL. CELL. BIOL. 15:102-111(1995).  
 CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE  
 CC RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE  
 CC ADENOVIRUS E4 GENE.  
 CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.  
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D13318; G286027; -  
 DR EMBL: U13044; G531893; -  
 DR MIM: 600609; -  
 DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
 DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
 DR PRAM: PF00178; Ets; 1.  
 DR HSP: 000422; IAWC.  
 DR TRANSFAC: T01390; -  
 KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DNA\_BIND 320 400 SS -> RC (IN REF. 2).  
 FT CONFLICT 289 290 A -> V (IN REF. 2).  
 FT CONFLICT 440 440  
 SQ SEQUENCE 454 AA: 51295 MW: 50E85D4 CRC32;

Query Match 16.2%; Score 289; DB 1; Length 454;  
 Best Local Similarity 27.8%; Pred. No. 2.8e-15;  
 Matches 85; Conservative 49; Mismatches 126; Indels 46; Gaps 9;

QY 60 LAFEDLPPDSSMAKAGASREPPPEPPCQVYDSQAPAGSLDVPGLTLEHSL 119  
 DB 108 LNTLETVKPADYVEVVIDDAHAESEALVEAQVI-----LLDCKHTITTSDETS 160  
 QY 120 EQVQSWVGEVLKDIETACKLNTADPMDSPSNVQWKLMTETQYKLPNGAFOELA 179  
 DB 161 EGYTRAA--ALEGYRKEOERLGIPIYDPLQWSTQVLMVYVWVKNFESMTDIDITLINIS 218  
 QY 180 GRELCAHSEDFRQSPPLGGDVLHAHLDIWKSAAWKKER----- 219  
 DB 219 GRELSLNQEDFFORVP--RGEITLMSHLELLRYYVLAQOQMNELVITDOVQIIPASVO 277  
 QY 219 -TSPGAIHCASTSEESMTD-----SEVDSQSGO-----PHLMQFLKELLKPSYSG 266  
 DB 278 SATPTTIKIVNSAKAAVQAPRISGEDRSSPGNRGTNNQIOLQWFLLELLDKARD 337  
 QY 267 REIRNLNKEGFIKIEDSAQVAFRLGIRKNRPANMYDKLSISQYKKGIIRKPDLSOR 326  
 DB 338 -CISWVGDE-GEFKLNGEELVAOKWQKRNKPTMYNTEKLSALNTYTDGDMICKVO-GKR 394  
 QY 327 LVYQGV 332  
 DB 395 FVYKPV 400

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE P135-GAG-MYB-ETS TRANSFORMING PROTEIN (FRAGMENT).  
 GN V-MYB AND V-ETS.  
 OS AVIAN LEUKEMIA VIRUS E26.  
 OC VIRUSES; RETROID VIRUSES; RETROVIRIDAE.  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 84068165.  
 RA NUNN M.F., SEEBURG P.H., MOSCOVICI C., DUESBERG P.H.;  
 RT "Tripartite structure of the avian erythroblastosis virus E26  
 RT transforming gene".  
 RL NATURE 306:391-395(1983).  
 CC -1- FUNCTION: MYB IS A DNA-BINDING PROTEIN THAT SPECIFICALLY RECOGNIZE  
 CC THE SEQUENCE YAA(G/T)G.  
 CC -1- DOMAIN: THE NEGATIVE REGULATORY DOMAIN AND THE C-TERMINAL DIFFER  
 CC CONSIDERABLY FROM THOSE OF OTHER MYB PROTEINS.  
 CC -1- DISEASE: MYB-ETS PROTEIN INDUCES PREDOMINANTLY ERYTHROBLASTOSIS IN  
 CC CHICKEN & TRANSFORMS AVIAN ERYTHROBLASTS & IMATURE MYELOMONOCYTIC  
 CC CELLS IN CULTURE. IT APPEARS THAT THE ETS DOMAIN IS RESPONSIBLE  
 CC FOR THE EFFECTS ON ERYTHROID CELLS AND THAT THE MYB DOMAIN ENCODES  
 CC THE MYELOID-TRANSFORMING CAPACITY.  
 CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X00144; G809070; -  
 DR PIR: B01348; TYVES.  
 DR PROSITE: PS00037; MYB\_1; PARTIAL.  
 DR PROSITE: PS00345; MYB\_2; PARTIAL.  
 DR PROSITE: PS00346; ETS\_DOMAIN\_1; 1.  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
 DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
 DR PRAM: PF00178; Ets; 1.  
 DR HSP: P14921; 2STW.  
 DR TRANSFAC: T01398; -  
 DR TRANSFAC: T01594; -  
 KW TRANSFORMING PROTEIN; POLYPROTEIN; DNA-BINDING; NUCLEAR PROTEIN.  
 FT NON\_TER 1 1  
 FT CHAIN 1 178 TRANSFORMING PROTEIN MYB.  
 FT CHAIN 179 669 TRANSFORMING PROTEIN ETS.  
 FT DOMAIN 90 142 TRANSCRIPTIONAL ACTIVATION DOMAIN  
 FT (BY SIMILARITY).  
 FT DNA\_BIND 556 640 ETS-DOMAIN.  
 FT SEQUENCE 669 AA: 74726 MW: 25DDB5A CRC32;

Query Match 16.1%; Score 288; DB 1; Length 669;  
 Best Local Similarity 23.98%; Pred. No. 5.0e-15;  
 Matches 121; Conservative 47; Mismatches 141; Indels 198; Gaps 21;

QY 15 HLLPPTVSTGTLEKAAAGVGLERDWSPPAPGEGLSAFYLSYPMLEPDESSNA 74  
 DB 143 HSTVADNTMTSG--DNAPVSLG--EHHHCIPSPV--DRHSTSEMAYMDTTGSGTPV 198  
 QY 75 AKAP-----GASSREPPPEP-----ECCPVIDSQA-----PAG-----SLDLPVPG 111  
 DB 199 LARPGVWGASSCCEDPWPMPQRLQASCCPPRSCCPWDEAALQEVPTGLEHSTDMECAD 258  
 QY 112 LLEHSLAQVQSWVGEVLKDIETACKLNTADPMDSPSNVQWKLMTETQYKLPNGA 171  
 DB 259 VPLLPSKKEWQALATKATSGFAKEDQRLGIPKDPQWTEHVRQWMAVAVNEFSL--K 316  
 QY 172 GKARQV--LAGKELCANSEDFRQSP--LGGDVLHAHLDIWK-----SAAMWK 216  
 DB 317 GVDYDFKCMNGAALCAIGKCEFLAPDFGVDGILMHEILQKEAKKPYANGVNAAPPE 376  
 QY 217 ER-TSPGAIH-----CASTSE-----ESW----- 236

[illegible]

```

RESULT 7
GABA_MOUSE
ID GABA_MOUSE STANDARD: PRT: 454 AA.
AC Q00422:
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GA BINDING PROTEIN ALPHA CHAIN (GABP-ALPHA SUBUNIT).
GN GABPA.
OS MUS MUSCULUS (MOUSE).
OC EUARCTOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MORIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91343912.
RA LAMARCO K., THOMPSON C.C., BYERS B.P., WALTON E.M., MCKNIGHT S.L.;
RT "Identification of Ets-and notch-related subunits in GA binding
RT protein."
RL SCIENCE 253:789-792(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 320-320.
RX MEDLINE: 98128030.
RA BATCHELOR A.H., PIPER D.E., DE LA BROUSSE F.C., MCKNIGHT S.L.,
RA WOLBERGER C.;
RT "The structure of GABPalpha/beta: an ETS domain-ankyrin repeat
RT heterodimer bound to DNA."
RL SCIENCE 279:1037-1041(1998).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
CC RICH REPEATS (GA REPEATS).
CC -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: UBICUITOUS.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: M74515; GI93383; -.
CC PIR: A40858; A40858.
CC PDB: 1AMC; 18-MAR-98.
CC MGD: MGI:95610; GABPA.
CC PROSITE: PS00345; ETS_DOMAIN_1; 1.
CC PROSITE: PS00346; ETS_DOMAIN_2; 1.
CC PROSITE: PS50061; ETS_DOMAIN_3; 1.
CC PFM: PF00178; Ets; 1.
CC TRANSFAC: T00298; -.
CC TRANSFAC: T01402; -.
CC TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN; 3D-STRUCTURE.
KW

```

	DNA_BIND		ETS-DOMAIN.
FT	320	400	
SEQUENCE	454 AA;	51363 MW;	C3B45F76 CRC32;

Query Match	15.5%	Score 277	DB 1:	Length 454
Best Local Similarity	28.8%	Pred. No.	2,4e-14	
Matches	83	Conservative	43	Mismatches 116; Indels 46; Gaps 9

[illegible]

RESULT 8

ID	FL11_XENLA	STANDARD;	PRT;	453 AA.
----	------------	-----------	------	---------

DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE RETROVIRAL INTEGRATION SITE PROTEIN FLI-1 HOMOLOG

OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
 OC EUKAROTA; METAZOA; CHORDATA; VERTEBRATA; BATRACHIA; ANURA;  
 OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.

RP SEQUENCE FROM N.A.  
RX MEDLINE; 94206844.  
RA MEYER D., WOLFF C.M., STIEGLER P., SENAN F., BEFORT N.,  
SECRETARY I. THUR

RT<sup>+</sup> XI-11, the *Xenopus* homologue of the *h11-1* gene, is expressed during RT embryogenesis in a restricted pattern evocative of neural crest cell distribution.

RL MECH. DEV. 44:109-121 (1993)

CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.

CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: X66979: G505487: -.  
DR PROSITE: PS00345: ETS\_DOMAIN\_1: 1.  
DR PROSITE: PS00346: ETS\_DOMAIN\_2: 1.  
DR PROSITE: PS50061: ETS\_DOMAIN\_3: 1.  
DR PFAM: PF00178: Ets: 1.  
DR HSSP: Q01543: 1FLI: -.  
DR TRANSFAC: T02067: -.  
DR TRANSCRIPTION REGULATION: ACTIVATOR: DNA-BINDING; NUCLEAR PROTEIN  
FT DNA\_BIND 282 362 ETS-DOMAIN.  
SQ SEQUENCE 453 AA: 51015 MW: DEB5A90B CRC32:

Query Match 15.2%; Score 272; DB 1; Length 485;  
Best Local Similarity 24.2%; Pred. No. 6,4e-14;  
Matches 112; Conservative 42; Mismatches 109; Indels 200; Gaps 18;

DR EMBL; M20508; G211464; JOINED.  
DR EMBL; M20509; G211464; JOINED.  
DR EMBL; M20510; G211464; JOINED.  
DR EMBL; M20511; G211464; JOINED.  
DR EMBL; M20512; G211464; JOINED.  
DR EMBL; M20513; G211464; JOINED.  
DR EMBL; M20514; G211464; JOINED.  
DR PIR; A28875; TVCHET.  
DR PROSITE; PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE; PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE; PS50061; ETS\_DOMAIN\_3; 1.  
DR PIR; P27577; ETS; 1.  
DR HSP; P27577; 1ETD.  
DR TRANSFAC; T00115; -.  
KW PROTO-ONCOGENE; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.  
FT DNA\_BIND 379 459 ETS-DOMAIN.  
FT CONFLICT 220 220 S -> C (IN REF. 2).  
FT CONFLICT 222 222 P -> R (IN REF. 2).  
FT FT 222  
SQ SEQUENCE 485 AA; 55021 MW; C58D8897 CRC32;

65 MLYPDESSMAAAP-----GASREPPPEP-----ECCPYIDQA-- 101  
2 MRYIDYITIGSGPPLPAPGVMOGASSCCEDPMPCRLASACCPRSCCPDEAIAOE 61  
101 -PAG---SLDLPVPGGLTLEEHSLEOVQSNVYGVKLDIETACKLINTADPMDWSPSY 155  
62 VPTGLEHYSSTDMECADVPLTPPSKEMMSQALKATFSGFAKEQORLOIPDPOOMTETHV 121  
156 QKWLLTETHOYRLPPKGAFOE--LAGKELCAMSEOFQRSP- LGGDVLAHLADIWK-- 211  
122 KQWVAWVAWNESSL- KGVDFQKFCMNAALCALKECELELADPDVDILMEHLEILOE 179  
211 -----SAAMKKER-TSPGAIHY-----CASTSE-----ESV----- 236  
180 EAKPYANGVAAPPESTYTSIDYISIGIHACQVPPSESEPSITESTIQILHPIISSEE 239  
236 -----TDS-----EV---DSSCSGQ--- 248  
240 LLSLTKENDYSVTLRDPVOTDSLDQIDYFTIKOEYVTPDMMCMGRASRGKLAGGDSFESI 299  
248 ----- 248  
300 ESYSDCDRLQSWSSQSSFSQLRVPYSDFSESDYPPALPNNKPKGTCKDYVRDRADNN 359  
248 -----PIHMOFLKELLAKHSIGRIPLNNKKGIFKEDSNOVAR 289  
360 KDRKVIIPAAALAGTGGGFIQWOLFLE-LTLDKSCQSFISWTG-DGWEFKLSDPDEVAR 417  
290 LMGIRKNRPAMANTDKLSRSIRQYKKGIIKPKDIPDSQRLAYOVFP 332  
418 RMGRKKKPKMNYEKLSRGLRYYIDKNIITHK-TAGRRYYRREV 459

RESULT 10  
FILE\_MOUSE ID FILE\_MOUSE STANDARD: PRT; 452 AA.  
AC P26323;  
DT 01-MAY-1992 (REL. 22, CREATED)  
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE RETROVIRAL INTEGRATION SITE PROTEIN FLI-1.  
DE FLI1 OR FLI-1.  
GN MUS MUSCULUS (MUSE).  
OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RC SEQUENCE FROM N.A.  
PC STRAIN-BALB/C; TISSUE-SPLEEN.

```

RX MEDLINE: 91257578.
RA BEN-DAVID Y., GIDDENS E.B., LETWIN K., BERNSTEIN A.;
RT "Erythroleukemia induction by Friend murine leukemia virus:
RT insertional activation of a new member of the ets gene family, FLI-1,
RT closely linked to c-ets-1."
RL GENES DEV. 5:908-918(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE: 93275657.
RA ZHANG L., LEVARCHANDEL V., ROMEO P.-H., BEN-DAVID Y., GREER P.,
RA BERNSTEIN A.;
RT "The flt-1 proto-oncogene, involved in erythroleukemia and Ewing's
RT sarcoma, encodes a transcriptional activator with DNA-binding
RT specificities distinct from other Ets family members."
RL ONCOGENE 8:1621-1630(1993).
CC -1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES
CC THE DNA SEQUENCE 5'C(CA)GGAGCT-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH HEMATOPOIETIC AND
CC NONHEMATOPOIETIC TISSUES.
CC -1- INDUCTION: SPECIFICALLY UP-REGULATED IN LEUKEMIC CLONES WITH
CC F-MULV INSERTIONS UP-STREAM OF THE FLI-1 LOCUS.
CC -1- DISEASE: INVOLVED IN ERYTHROLEUKEMIA INDUCTION BY FRIEND MURINE
CC LEUKEMIA VIRUS (F-MULV).
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch).
CC -----
DR EMBL: X59421: G50975: -
DR PIR: S17403: S17403.
DR MGI: 95554: FLI1.
DR PROSITE: PS00345: ETS_DOMAIN_1: 1.
DR PROSITE: PS00346: ETS_DOMAIN_2: 1.
DR PROSITE: PS50061: ETS_DOMAIN_3: 1.
DR PFM: PF00178: Ets; 1.
DR HSP: Q01543: FLI1.
DR TRANSFAC: T01408: -
RW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;
KW PROTO-ONCOGENE.
FT DNA_BIND 281 361 ETS-DOMAIN.
SQ SEQUENCE 452 AA; 51002 MW; CECACDD4 CRC32;

Query Match 15 2% Score 272; DB 1; Length 452;
Best Local Similarity 25.6% Pred No. 5.8e-14;
Matches 102; Conservative 46; Mismatches 142; Indels 108; Gaps 16;

OY 2 GSASPOLSSVSPSHLLPPTVSRITGLEKAAAGVLEERRDMSPPSPATPEOG---LSA 57
OY 3 GTIKELSLVSPDQSL-----FDAYGAHAHLPRADMTAS--GSPDVGOPHKINP 50
OY 58 FLTSLFDMLIPEDSSWAAPGASNEF-----PPEEDQCPVIDSQAPAGSLDIVPGG 111
OY 51 L-----PQOEWINPVRVNRKREYDHNMGSRESFVDCSVSKNKLVGCGEAMPNN 101
OY 112 L--LEHSLLEQVSMVVEVLKDIETACKLINTADPMDSPSNVOKLMTENHOYRLP 169
OY 102 VNSYDEKNGPPPPNNTTTE-----RRVIVPADPTLMTQEHVROWLEVAIKEYGLM 152
OY 170 PMGKA-FOELACKELCAMEEQF-RORSPPLGVDLFAHL----- 207
OY 153 EIDTSFQAMMDKELCKMKKEDPLRTASVNTFVLSHLSYLRSSSLAYNTTSHTDOSS 212
OY 207 -----DIWKSAM-----MKERTSPGAIHYCASTSEE 233
OY 213 RLNVKEDPSYDVSVRGAMNNMNSGLNKSPLLGGSTQTMCKNTQORPODPYQILGPTS-- 271

```

```

OY 234 SMTSEVDSSCGOPILHMOFLKELLKPSYGRIRBLNKEKFIKIEDSAOYARLWGI 293
OY 271 ----SRLNPSGGQ-IQMOFLLE-LSDSASNCITV-ESTNGEFKMTDPDEVARRWGE 323
OY 294 RKNRPAMNYDLRSIRIOYKRGIIKRPDISORLYVOP 331
OY 324 RSKRPNMYDKLSRALRYRYKNIMTYH--GKRAYKRF 360
DB

RESULT 11
FLI1_HUMAN STANDARD; PRT; 452 AA.
ID FLI1_HUMAN
AC 001543:
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FLI-1 ONCOGENE (ERGB TRANSCRIPTION FACTOR).
GN FLI1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BONE MARROW;
RX MEDLINE: 92396239.
RA DELATRE O., ZUCMAN J., PLOUGASTEL B., DESMAZE C., MELOT T.,
RA PETER M., KOVAR H., JOUBERT I., DE JONG P., ROULEAU G.;
RT "Gene fusion with an ETS DNA-binding domain caused by chromosome
RT translocation in human tumours."
RL NATURE 359:162-165(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93075640.
RA WATSON D.K., SMITH F.E., THOMPSON D.M., CHENG J.Q., TESTA J.R.,
RA PAPAS T.S., SEITZ A.;
RT "The ERGB/FLI-1 gene: isolation and characterization of a new member
RT of the family of human ETS transcription factors."
RL CELL GROWTH DIFFER. 3:705-713(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX PRASAD D.D., RAO V.N., REDDY E.S.;
RT "Structure and expression of human FLI-1 gene."
RL CANCER RES. 52:5833-5837(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RX MEDLINE: 93176799.
RA HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MAKI R.A.,
RA BECK E., KLEMSZ M.J.;
RT "Human FLI-1 localizes to chromosome 11Q24 and has an aberrant
RT transcript in neuroepithelioma."
RL BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).
RN [5]
RP STRUCTURE BY NMR OF 276-373.
RX MEDLINE: 95292091.
RA LIANG H., MAO X., OLEJNICZAK E.T., NETTESHEIM D.G., YU L.,
RA MENDONS R.P., THOMPSON C.B., FESIK S.W.;
RT "Solution structure of the ets domain of FLI-1 when bound to DNA."
RL NMR. STRUCT. BIOL. 1:871-875(1994).
CC -1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES
CC THE DNA SEQUENCE 5'C(CA)GGAGCT-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DISEASE: A FORM OF EWING'S SARCOMA IS CHARACTERIZED BY A
CC CHROMOSOMAL TRANSLOCATION T(11;22)(Q24;Q12) WHICH INVOLVES FLI1
CC AND EWS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X67001; G32530; -  
 DR EMBL: M98833; G182189; -  
 DR EMBL: S45205; G257354; -  
 DR EMBL: M93255; G182660; -  
 DR PDB: 1FLI; 15-SEP-95.  
 DR MIM: 193067; -  
 DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
 DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
 DR PFAM: PF00178; Ets; 1.  
 DR TRANSFAC: T02066; -  
 DR TRANSSCRIPTION REGULATION: ACTIVATOR: DNA-BINDING; NUCLEAR PROTEIN;  
 KW PROTO-ONCOGENE: CHROMOSOMAL TRANSLOCATION; 3D-STRUCTURE.  
 FT DNA\_BIND 281 361  
 FT CONFLICT 69 69 ETS\_DOMAIN.  
 FT CONFLICT 77 77 MISSING (IN REF. 4).  
 FT CONFLICT 130 130 P -> A (IN REF. 4).  
 FT CONFLICT 133 133 W -> V (IN REF. 4).  
 FT CONFLICT 323 323 E -> Q (IN REF. 3 AND 4).  
 FT CONFLICT 391 391 Y -> Q (IN REF. 2).  
 FT CONFLICT 426 426 MISSING (IN REF. 2 AND 4).  
 SO SEQUENCE 452 AA: 50982 MM: 19247677 CRC32;

Query Match 15.1%; Score 270; DB 1; Length 452;  
 Best Local Similarity 25.9%; Pred. No. 8.3e-14;  
 Matches 103; Conservative 50; Mismatches 137; Indels 108; Gaps 17;

QY 2 GSASPGLSVSPSHLLPDDTVSRITGLERKAAGAVGLERDMSPPPATPEOG-----LSA 57  
 DB 3 GTIKELALSVSDQSL-----FDSAYGAHAHLPRADMTAS--GSPDYGPCHKINP 50  
 QY 58 FYLSYFDMLYPEDSSWAAPAGASREE-----PPEEPCQPVDSOAPAGSLDLYPGG 111  
 DB 51 L-----PQOEWINQPVRYNVRREYDHNGSRESYVDCSV-----SICSLVYGG 95  
 QY 112 LT-----LEHSLLEQVQSMVGVGLKDITACKLNTADPMDSPPSNVQKWLMT 163  
 DB 96 ESNPMYNSYMDKNGKPPPPMTTNE-----RRVYPADPPLMTQEHVROWLEWAI 146  
 QY 164 HOYRLPPMGKA-FOELAGKELCAMSEEPORSPPL-GGDVLAHL----- 207  
 DB 147 KEYSLEIDTSFQNMNDGKELCKNKEDPLRATITLYNTEVLSHLSYRESLLAYNTTS 206  
 QY 207 -----DLTKSAAMKERTSP-----GAIHYCASTSEE----- 234  
 DB 207 HTDQSSRLSVKEDPSYDSVRRGAMGNMNSGLNKSPPLGAGOTISKNTGQRPDPYOIL 266  
 QY 234 SWTDEVDSSCGOPHILMOFLKELLKPHSGRIIRLNKEGFKIEDSAOVARLMGI 293  
 DB 267 GPTSSRLANPGSGQ-IQLMOPLE-LLSDSANASCITW-EGTNEEFKMTDPDEVARRMGE 323  
 QY 294 KKNRPAMNYDKLSRSIROYYKGIIRKPDISORLYQYF 331  
 DB 324 RKSKEPMNYDKLSRALRYLYDKNIMTKVH-GKRAYAKF 360

RESULT 12  
 ETV6\_MOUSE  
 ID ETV6\_MOUSE STANDARD: PRT: 485 AA.  
 AC P97360;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE ETS-RELATED PROTEIN TEL (ETS TRANSLOCATION VARIANT 6).  
 GN ETV6 OR TEL.  
 OS MOS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC ROSENTIA; SCIURGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J;  
 RA BERNARD O.;

RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y07915; E274398; -  
 DR MGD: MGI.109336; ETV6  
 DR PROSITE: PS00345; ETS\_DOMAIN\_1; FALSE\_NEG.  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
 DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
 DR PFAM: PF00178; Ets; 1.  
 DR HSSP: Q01543; 1FLI.  
 DR NUCLEAR PROTEIN; DNA-BINDING. ETS-DOMAIN.  
 FT DNA\_BIND 335 416  
 SO SEQUENCE 485 AA: 56405 MM: C1A/5087 CRC32;

Query Match 14.9%; Score 265.5; DB 1; Length 485;  
 Best Local Similarity 21.9%; Pred. No. 2e-13;  
 Matches 98; Conservative 47; Mismatches 80; Indels 223; Gaps 15;

QY 75 AKAPGASREE-----PPEEPCQPVDSOAPAGSLD--VPGILTEHSLQVQSMV 126  
 DB 2 SETPAGSSIKQEIISTTPESP-----VASHRSSTPLHYHTVPRALRREDSIH----- 51  
 QY 127 VGEVLADITACKLNTADPMDSPPSNVQKWLMTTEHOYRLPPMGKAFOLAGKELCAM 186  
 DB 51 -----LPTHLRL-----QPIYMSRDVAVQWLKNAENFSLRPLESKKFEKNGKALLIL 98  
 QY 187 SEQFROBSPPLGGDYHL-----AHLDTKSAAMKERTS-----PG----- 223  
 DB 99 TKEDFYRSPHSGDYLYELLOHI-----LKQKSHMLESPEEPEDSIHTKPEVLH 150  
 QY 223 -----AIHYCASTSE-----ESWTS----- 233  
 DB 151 QNHEDNCVQRTPRPAEYVHNHPTIELHRPSPITTNRPSPDEPQORPQSPIDNM 210  
 QY 233 ----- 239  
 DB 211 SRLSPVERKAGPRLQOENNHQETPLSVSPVENNHCLPSSPMQESTRYQLMSPIMHP 270  
 QY 239 -----EVDSSCS-----GQPIH----- 251  
 DB 271 LLINRHSHSVDFKQSRSEDOANREGKPIVLSHREDLAYLNIIVMSPEEHAMPGR 330  
 QY 251 -----LMOFLKELLKPHSYGRIIRLNKEGFKIEDSAOVARLMGIRKPNRPMYDK 304  
 DB 331 IADCRLLMDYVYQ-LLSDSYRENFIRMEDKSKIFRIYDPNGRLARLMGNKRNKRNMTYK 389  
 QY 305 LRSIROYYKGIIRKPDISORLYQYF 332  
 DB 390 MSRALRHYKLNILRK-EPGRLLEFRM 416

RESULT 13  
 ETV6\_HUMAN  
 ID ETV6\_HUMAN STANDARD: PRT: 452 AA.  
 AC P41212;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE ETS-RELATED PROTEIN TEL (ETS TRANSLOCATION VARIANT 6).  
 GN ETV6 OR TEL.





CC EXTREMITY: P54 AND P68.  
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M22462; G211753; -  
 DR EMBL: X13026; G63180; -  
 DR EMBL: X13027; G63383; -  
 DR PIR: A31285; TVCHTE;  
 DR PIR: S07625; S07625;  
 DR PIR: S29132; S29132;  
 DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1;  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1;  
 DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1;  
 DR PRAM: PF00178; Ets; 1;  
 DR HSP: P27577; LETD;  
 DR TRANSFAC: T00114; -  
 KM PROTO-ONCOGENE: DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING;  
 KM PHOSPHORYLATION;  
 FT DOMAIN 34 135 POINTED.  
 FT DNA\_BIND 335 415 ETS\_DOMAIN.  
 FT MOD\_RES 38 38 PHOSPHORYLATION (BY MAPK) (BY  
 FT SIMILARITY)  
 FT CONFLICT 117 117 R -> A (1N REF. 3).  
 FT SEQUENCE 441 AA: 50326 MW: 734E39B3 CRC32;

Query Match 14.1%; Score 252.5; DB 1; Length 441;  
 Best Local Similarity 24.0%; Pred. No. 1.8e-12;  
 Matches 96; Conservative 38; Mismatches 89; Indels 177; Gaps 15;

QY 105 LDLPV-----GGTLLEHSLDVOGVMVGEVLKDIETACKLINTADPMDSPSNVQW 158  
 DB 21 IDLFSPDMECADVPLTPSSKEMMSQALKATFSGFAKQRLGIPKDPQWTEHVDM 80  
 QY 159 LMTHEQVRLPPMGKAFQF--LAKELCAMSEDFRQSP--LGSDVLAHLIDWK----- 211  
 DB 81 VMVAVVEFSL--KGVDFOFCMNGAALCALGKCEFLERPRDVGILWEHLLEILQKEPAK 138  
 QY 211 -----SAAMKERTSPGAIRH-----CASTSE-----ESW----- 236  
 DB 139 PTPANGVNAAYPESRTYSDFISYGIENACVPPSEFSEPSFTESYOTLHPISSELLS 198  
 QY 236 -----TDS-----EV-----DSSCSGO----- 248  
 DB 199 LKENDYPSVILRDVPVQDSLOTQDTFTKQEVYVPDNCMKGRASRGILGODSESIESTY 258  
 QY 248 ----- 248  
 DB 259 DSCRLTOSWSSQSFOSLQRPVPSYDFSEDYPALPNKPKGTFFDYVDRADNMKDK 318  
 QY 248 -----PIHLMOFLKELLKPHSYGRTIRMLNKEGIRKIEDSAQVAPRLMG 292  
 DB 319 PVIPAAALAGTGSGLPILMOFLLE-LLTDKSCOSFISWVG-DQWERKLDPDDEVARRMG 376  
 QY 293 IRKNRPANVYKLSRSIRQYKKGIRKRPDISORLAVYQFV 332  
 DB 377 KRKNPKMNYEKLRSGLRYYYDKNIHR-TAGKRYVYRFV 415

RESULT 15  
 ID ETSB\_XENLA STANDARD: PRT: 472 AA.  
 AC 091712;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE C-ETS-2B PROTEIN.  
 GN ETS2B OR ETS-2B  
 OS XENOPUS LAEVIS (AFRICAN CLAMLED FROG).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
 NC MESOBATRACHIA; PIPOIDEA; PIPOIDAE; PIPOIDAE; XENODIDINAE; XENOPUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARY;  
 RX MEDLINE: 90356411.  
 RA WOLFF C.M., STIGLER P., BALTZINGER M., MEYER D., GHYSDAEL J.,  
 RT STEHELIN D., BEFORT N., REMY P.;  
 RT "Isolation of two different c-ets-2 proto-oncogenes in Xenopus  
 RT laevis.";  
 RL NUCLEIC ACIDS RES. 18:4603-4604(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-OVARY;  
 RX MEDLINE: 92088972.  
 RA WOLFF C.M., STIGLER P., BALTZINGER M., MEYER D., GHYSDAEL J.,  
 RA STEHELIN D., BEFORT N., REMY P.;  
 RT "Cloning, sequencing, and expression of two Xenopus laevis c-ets-2  
 RT protooncogenes.";  
 RL CELL GROWTH DIFFER. 2:447-456(1991).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X52635; G64621; -  
 DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1;  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1;  
 DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1;  
 DR PRAM: PF00178; Ets; 1;  
 DR HSP: P14921; ZSTW.  
 KM DNA-BINDING; NUCLEAR PROTEIN.  
 FT DNA\_BIND 366 446 ETS\_DOMAIN.  
 FT SEQUENCE 472 AA: 53928 MW: 196506B4 CRC32;

Query Match 14.0%; Score 250; DB 1; Length 472;  
 Best Local Similarity 22.9%; Pred. No. 3.1e-12;  
 Matches 92; Conservative 35; Mismatches 91; Indels 184; Gaps 15;

QY 108 VPGLTLEHSL-----LEQVQSMVWGEVLKDIETACKLINTADPMDSPSN 154  
 DB 52 VPTGLDSYHDSSECLPLTPPCSKAVVSAQKIDTFNCFAKERQ-LGIPNPVLMWENN 110  
 QY 155 VOKMLTTEHQVRLPPMGKAFQF--LAKELCAMSEDFRQSP--LGSDVLAHLIDWK----- 208  
 DB 111 VQWLLMAKERSLENVN--FOKFLMNGHELCSLGERFLALADPFGDILMELEEMAK 168  
 QY 208 -----IMKSAAMKERTSP----- 222  
 DB 169 EYQKAKQEPYIDHSNRDLSLNMNANDSLNTFADPLQCGAOVHNPKNMGVNDMCSVPTGQ 228  
 QY 222 -----GAIHVCASTSE----- 233  
 DB 229 TLNPKQEFQOYPSCLSKSRVAVNPYPASQDPARSHMNVLLNSLNGRLRYDSDGSGTES 288  
 QY 233 -----ESWTDSE-----VDSSC----- 245  
 DB 289 FESTESLHSMTSOSSLYDMQRPVSYDSFEEDGNQTLCKKQPMSEFDYIÖDCEPAELG 348  
 QY 245 -----SGQPIHLMOFLKELLKPHSYGRTIRMLNKEGIRKIEDSAQVAPRLMG 290  
 DB 349 KPVIPASTIAGTGSGLPILMOFLLE-LLTDKSCOSFISWVG-DQWERKLDPDDEVARR 405



RESULT 2  
ERG\_HUMAN STANDARD: PRT: 462 AA.  
ID ERG\_HUMAN  
AC P11308;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE TRANSCRIPTIONAL REGULATOR ERG (TRANSFORMING PROTEIN ERG) (ERG-2/ERG-1).  
GN ERG.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
RN PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RP [1]  
RX MEDLINE: 87263429.  
RA RAO V.N., PAPAS T.S., SHYAM E., REDDY P.;  
RT "erg, a human ets-related gene on chromosome 21: alternative  
RT splicing, polyadenylation, and translation.";  
RL SCIENCE 237:635-639(1987).  
RN [2]  
RP SEQUENCE OF 100-462 FROM N.A.  
RX MEDLINE: 87317608.  
RA REDDY E.S.P., RAO V.N., PAPAS T.S.;  
RT "The erg gene: a human gene related to the ets oncogene.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 84:6131-6135(1987).  
RN [3]  
RP CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE: 94356859.  
RA DUINN T., PRAISSMAN L., HAGAG N., VIOLA M.V.;  
RT "ERG gene is translocated in an Ewing's sarcoma cell line.";  
RL CANCER GENET. CYTOGENET. 76:19-22(1994).  
RN [4]  
RP CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE: 94243799.  
RA ICHIKAWA H., SHIMIZU K., HAYASHI Y., OHKI M.;  
RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human  
RL myeloid leukemia with t(16;21) chromosomal translocation.";  
CC CANCER RES. 54:2865-2868(1994).  
CC [1]- FUNCTION: ACTS AS A TRANSCRIPTIONAL ACTIVATOR (BY SIMILARITY).  
CC [1]- SUBCELLULAR LOCATION: NUCLEAR.  
CC [1]- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED  
CC BY A CHROMOSOMAL TRANSLOCATION T(16;21)(P11;Q22) WHICH INVOLVES  
CC ERG AND FUS.  
CC [1]- SIMILARITY: BELONGS TO THE ETS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch)).  
CC -----  
CC EMBL: M17254, G182187; -;  
CC EMBL: M21535, G182185; -;  
CC PIR: A29515; TVHUEG.  
CC MIM: 165080; -;  
CC DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
CC DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
CC DR PROSITE: PS50061; ETS\_DOMAIN\_3; 1.  
CC DR PFAM: PF00178; Ets; 1.  
CC DR HSP: Q01543; IFLI.  
CC DR TRANSFAC: T00265; -;  
CC DR TRANSFAC: T00266; -;  
CC KW TRANSCRIPTION REGULATION; ACTIVATOR; NUCLEAR PROTEIN; DNA-BINDING;  
CC KW TRANSFORMING PROTEIN; ALTERNATIVE INITIATION; PROTO-ONCOGENE;  
CC KW CHROMOSOMAL TRANSLOCATION.  
CC FT CHAIN 1 462 TRANSFORMING PROTEIN ERG-2.  
CC FT DNA BIND 100 462 TRANSFORMING PROTEIN ERG-1.  
CC FT DNA BIND 294 374 ETS-DOMAIN.  
CC SQ SEQUENCE 462 AA; 52031 MW; CECC2DB CRC32;

Query Match 16.88; Score 300.5; DB 1; Length 462;  
Best Local Similarity 26.78; Pred. No. 3,7e-16;  
Matches 103; Conservative 48; Mismatches 118; Indels 117; Gaps 17;  
OY 45 PSPATPEQGLSAF-----YLSYFDM-----YEDSSNAKAPGASSRE 84  
DB 6 PDPAHIREALSYSEDSLEFCAYGTPHLAKTEMASSSSSDGGQTSKMSPRVPODWLS 65  
OY 85 EPP-----EEEDQCPVIDSQAPAGSLDLP---GGLTFEHSLEQV 123  
DB 66 QPAPAVTIKMECNPSQVNGSRNSPDECSYAKGGMNOSPVTGMNNGSYEEHNM--PP 123  
OY 124 SMVYGEVLKDIETACKLINTADPMSPSNVQKWLMTHEYRLPPMG-KAQOELAGE 182  
DB 124 NATNE-----KRVYPADEPTLMSTDHVRQWLEMAVKEGLDVLNILLFONIDKE 174  
OY 183 LCAENSEDFRQSP-LGQVLAHL-----DIWKS----- 212  
DB 175 LCKMTKDDFOQLPSTYNADILLSHLHRTPLPHLSDVAKALONSRLMHAARTDLP 234  
OY 212 -----AAW-----MKERTSPGAIHYCASTSEE-----SWTSEVDSSCS 245  
DB 235 YEPPRRSAWKGHPPTQSKAPSPSTV---PKTEDQRPOLDPYQILGPTSSRLANPS 291  
OY 246 GQPIHLMQFLKELLKPSYGRFRLNKEKGFIEDSAQVARKLGNRPANNYDKL 305  
DB 292 GQ-IQLWQFLLE-LLDSSNSSCITW-BGTNGEFMTDDEVARRGERSKPNMNYDKL 348  
OY 306 SRSIRQYKAGITRKPDISQRLYQF 331  
DB 349 SRALRYIDKIMTKVH-GRRYAYKF 373  
RESULT 3  
ERG\_DROME STANDARD: PRT: 464 AA.  
ID ERG\_DROME  
AC Q04688;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE DNA-BINDING PROTEIN D-ELG.  
GN ETS97D OR ELG.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDEA;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93096481.  
RA THE S.M., XIE X., SMYTH F., PAPAS T.S., WATSON D.K., SCHULTZ R.A.;  
RT "Molecular characterization and structural organization of D-elg, an  
RT ets proto-oncogene-related gene of Drosophila.";  
RL ONCOGENE 7:2471-2478(1992).  
RN [2]  
RP SEQUENCE OF 323-463 FROM N.A.  
RX MEDLINE: 91319397.  
RA PRIEYL L.J., WATSON D.K., SCHULTZ R.A., PAPAS T.S.;  
RT "D-elg, a member of the Drosophila ets gene family: sequence,  
RT expression and evolutionary comparison.";  
RL ONCOGENE 6:1175-1183(1991).  
RN [3]  
RP SEQUENCE OF 298-449 FROM N.A.  
RX MEDLINE: 92249640.  
RA CHEN T., BUNTING M., KARIM F.D., THUMMEL C.S.;  
RT "Isolation and characterization of five Drosophila genes that encode  
RT an ets-related DNA binding domain.";  
RL DEV. BIOL. 151:176-191(1992).  
CC [1]- SUBCELLULAR LOCATION: NUCLEAR.  
CC [1]- SIMILARITY: BELONGS TO THE ETS FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



```

RP SEQUENCE FROM N.A.
RX MEDLINE: 93180783.
RA WATANABE H., SAWADA J.-I., YANO K.-I., YAMAGUCHI K., GOTO M.,
RT "CDNA cloning of transcription factor E4TF1 subunits with Ets and
RL notch motifs."
RN MOL. CELL. BIOL. 13:1385-1391(1993).
RP SEQUENCE FROM N.A.
RX MEDLINE: 95097980.
RA GUGNELA S., VIRBASIS J.V., SCARPULLA R.C.;
RT "Four structurally distinct, non-DNA-binding subunits of human
RT nuclear respiratory factor 2 share a conserved transcriptional
RT activation domain."
RL MOL. CELL. BIOL. 15:102-111(1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
CC RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE
CC ADENOVIRUS E4 GENE.
CC -1- SUBUNIT: HETEROTRIMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13318; G286027; -
DR EMBL: U13044; G51893; -
DR MIM: 600609; -
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR PFM: PF00178; Ets; 1.
DR HSSP: Q00422; IAWC.
DR TRANSFAC: T01390; -
DR TRANSCRIPTION REGULATION: DNA-BINDING; NUCLEAR PROTEIN.
FT DNA_BIND 320 400
FT CONFLICT 289 290 SS-> RC (IN REF. 2).
FT CONFLICT 440 440 A-> V (IN REF. 2).
SQ SEQUENCE 454 AA; 51295 MW; E0EA85D4 CRC32.

Query Match 16.24; Score 289; DB 1; Length 454;
Best Local Similarity 27.86; Pred. No. 2,8e-15;
Matches 85; Conservative 49; Mismatches 126; Indels 46; Gaps 9;

```

```

RESULT 6
MYB_AVILE ID MYB_AVILE STANDARD; PRT; 669 AA.
AC POL105;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE P135-GAG-MYB-ETS TRANSFORMING PROTEIN (FRAGMENT).
OS V-MYB AND V-ETS.
GN AVIAN LEUKEMIA VIRUS E26.
OC VIRUSES; RETROVIRUSES; RETROVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 84068165.
RA NUNN M.F., SEEBURG P.H., MOSCOVICI C., DUESBERG P.H.;
RT "Tripartite structure of the avian erythroblastosis virus E26
RT transforming gene."
RL NATURE 306:391-395(1983).
CC -1- FUNCTION: MYB IS A DNA-BINDING PROTEIN THAT SPECIFICALLY RECOGNIZE
CC THE SEQUENCE YAAC(G/T)G.
CC -1- DOMAIN: THE NEGATIVE REGULATORY DOMAIN AND THE C-TERMINAL DIFFER
CC CONSIDERABLY FROM THOSE OF OTHER MYB PROTEINS.
CC -1- DISEASE: MYB-ETS PROTEIN INDUCES PREDOMINANTLY ERYTHROBLASTOSIS IN
CC CHICKEN & TRANSFORMS AVIAN ERYTHROBLASTS & IMMATURE MYELOMONOCYTIC
CC CELLS IN CULTURE. IT APPEARS THAT THE ETS DOMAIN IS RESPONSIBLE
CC FOR THE EFFECTS ON ERYTHROID CELLS AND THAT THE MYB DOMAIN ENCODES
CC THE MYELOID-TRANSFORMING CAPACITY.
CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X00144; G809070; -
DR PIR: B01348; TYVES.
DR PROSITE: PS00337; MYB_1; PARTIAL.
DR PROSITE: PS00334; MYB_2; PARTIAL.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR PFM: PF00178; Ets; 1.
DR HSSP: P14921; 2STW.
DR TRANSFAC: T01398; -
DR TRANSCRIPTION REGULATION: POLYPROTEIN; DNA-BINDING; NUCLEAR PROTEIN.
FT NON_TER 1 178
FT CHAIN 179 669 TRANSFORMING PROTEIN MYB.
FT CHAIN 179 669 TRANSFORMING PROTEIN ETS.
FT DOMAIN 90 142 TRANSCRIPTIONAL ACTIVATION DOMAIN
FT DOMAIN 90 142 (BY SIMILARITY).
FT DNA_BIND 556 640
FT SEQUENCE 669 AA; 74726 MW; 25DD6B5A CRC32.

Query Match 16.18; Score 288; DB 1; Length 669;
Best Local Similarity 23.98; Pred. No. 5,6e-15;
Matches 121; Conservative 47; Mismatches 141; Indels 198; Gaps 21;

```

```

DB 259 VLLPSSKEMSQALKATFSGFAKQQLGIPKDPQQTETHYRWVWMAVNEFSL--K 316
QY 172 GAFADE--LAGELCAMSEOFROKSP-LGQVLAHNDIMK-----SAAMK 216
DB 317 GVDFOKFCNGAALCALGECLELAPDVGDLMEHLILQEKAPKPVANGVNAAYPE 376
QY 217 ER-TSPGAIHY-----CASTSE-----ESW----- 236
DB 377 SKRTSDYFTSYGIEHAQCVPESEFSEPTSTESTOTLHPISSEBELSLKENDYPSVILR 436
QY 236 ---TDS-----EY-----DSCSCGO----- 248
DB 437 DVGQDLSLDYDTFTIKQEVVTPDNMGKRVSGKLGQDPSFESIESYDSCDLTQSMSSQ 496
QY 248 ----- 248
DB 497 SSFOSLQRPVSYSDFSESDYPALPNNKPKGTFKDYVRADRNKDKPVIPAAALAGYTG 556
QY 248 --PIHLMQFLKELLKPHSYGRIRLWLNKKGIFKIEDSAOVARLWGIKRNPRAMNYDKL 305
DB 557 SGPLOKMOFLLE-LTQKSCOSTISWTG-DGWEFKLSDDPEVAKRWGKKRKNPKMDYEXL 614
QY 306 SRSIROYYKKGIIRKPDISOQLVYQFV 332
DB 615 SRGLRYDYDKNVIRK-TACKRYVYRFV 640

```

RESULT 7  
GABA\_MOUSE STANDARD: PRT: 454 AA.

```

AC 000422:
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GA BINDING PROTEIN ALPHA CHAIN (GABP-ALPHA SUBUNIT).
GN GABP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91343912.
RA LAMARCO K., THOMPSON C.C., BYERS B.P., WALTON E.M., MCKNIGHT S.L.;
RT "Identification of Ets- and notch-related subunits in GA binding
RT protein."
RL SCIENCE 253:789-792(1991).
RN [2]
RX X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 320-320.
RA MEDLINE: 98128030.
RA BATHELOIR A.H., PIPER D.E., DE LA BROUSSE F.C., MCKNIGHT S.L.,
RA WOLBERGER C.;
RT "The structure of GABPalpha/beta: an Ets domain-ankyrin repeat
RT heterodimer bound to DNA."
RL SCIENCE 279:1037-1041(1998).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
CC RICH REPEATS (GA REPEATS).
CC -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: UBIDUOUS.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74515; G193383;
DR PIR: A40858; A40858.
DR PDB: 1AMC; 1B-MAR-98.

```

```

DR MCD: MGI:95610; GABA.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS00361; ETS_DOMAIN_3; 1.
DR PFM: PF00178; Ets; 1.
DR TRANSFAC: T00298;
DR TRANSFAC: T01402;
KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN; 3D-STRUCTURE.
FT DNA_BIND 320 400 ETS-DOMAIN
SQ SEQUENCE 454 AA; 51363 MW; C3B45E76 CRC32;

```

Query Match 15.5%; Score 277; DB 1; Length 454;  
Best Local Similarity 28.8%; Pred. No. 2,4e-14;  
Matches 83; Conservative 43; Mismatches 116; Indels 46; Gaps 9;

```

QY 78 PGASSREPEPEEQCPYIDSOAPASLDLVPGLTLEHSLSDVOQSVNGEVLKDIETA 137
DB 126 PDHHAELAEHLVEEAQYI-----TLDGTRKHITTSIDETSEQYTRMAA--ALEGYRKE 176
QY 138 CILLNITADPMDWSPSNQKLLWTEHOYRLPRMKAFOELAGKELCAMSEOFROKSP 197
DB 177 CERLGIPIYDPIKWSIDQVLAHWVWVKESMTIDTLTINISGRLESLNDEDFORVP- 236
QY 198 GGDVLAHNDIKWSAAMKKE-----TSPGAIHYCASTS----- 232
DB 236 KGEILMHSLELLRKTYVLAQEQQNEIYTIQPVQIIRASVPATPTTIKYINSSAKAK 295
QY 232 -ESWTDEVDSSCSGO-----PIHLMQFLKELLKPHSYGRIRLWLNKKGIFKIEDS 284
DB 296 VQSRPISGEDRSSPGNRGTNGNGQIQLOMFLLELTDRDARD-CISWQDE-GEFKLNQ 353
QY 285 AOVARLWGIKRNPRAMNTDKLSRSIROYYKKGIIRKPDISOQLVYQFV 332
DB 334 ELVAQKQGRKPKPMYNEKISRALRYDDGMICKVO-GKREYKRV 400

```

## RESULT 8

FL11\_XENLA STANDARD: PRT: 453 AA.

```

AC P41157:
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE RETROVIRAL INTEGRATION SITE PROTEIN FL1-1 HOMOLOG.
GN FL1.
OS XENOPUS LAEVIS (AFRICAN CLAMWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94206844.
RA MEYER D., WOLFF C.M., STIEGLER P., SENAN F., BEFORT N.,
RA BEFORT J.J., REMY P.;
RT "X1-fl, the xenopus homologue of the fl1-1 gene, is expressed during
RT embryogenesis in a restricted pattern evocative of neural crest cell
RT distribution."
RL MECH. DEV. 44:109-121(1993).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING EMBRYOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X66979; G505487;
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.

```

DR PROSITE; PS50061; ETS\_DOMAIN\_3; 1.  
 DR PFAM; PF00178; Ets; 1.  
 DR HSSP; 001543; 1FLI.  
 DR TRANSFAC; T02067; .  
 KW TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN.  
 FT DNA\_BIND 282 362 ETS-DOMAIN.  
 SQ SEQUENCE 453 AA; 51015 MW; DEB5A90B CRC32;

Query Match 15.48; Score 275; DB 1; Length 453;  
 Best Local Similarity 32.48; Pred. No. 3,4e-14;  
 Matches 78; Conservative 29; Mismatches 78; Indels 56; Gaps 9;

QY 143 ITADPMDSNVOKMLWTEHOYRLPPMG-KAFQELAGKELCAMSEDF-RQRLSGD 200  
 DB 125 VPADPALMSDQHVQWEMAIKEGLVEIDCSLFQNLIDGELCKMKEDFLNSTSYNTE 184  
 QY 201 VLAHLIDKMSAA- 222  
 DB 185 VLLSHLYLDDSSSLGNTQAHTDQSSRLTAKEDPYEARRRSGMNSSPVTKSPM 244  
 QY 222 GAHYCASTSEESW- -TDSEVDSGSGOPTILMQLKELKPHSYGRFR 270  
 DB 245 GGNQNVKSSDQSQRPDIQIIGPTISSRLANPGSG-QIQLWQFLE-LSDSSNASCIT 302  
 QY 271 WLKKEGIFEDSAOVARLWIRKNRPANNDKLSRSIROYKKGIIKRPDISQRLVYO 330  
 DB 303 W-DGTNGEFKMTDPDEVARMRGERSKPKNNNDKLSRALRYIDKSIKMHV-GRRYAK 360  
 QY 331 F 331  
 DB 361 F 361

RESULT 9  
 ETSB\_CHICK STANDARD; PRT; 485 AA.  
 AC P15062;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE TRANSFORMING PROTEIN P68/C-ETS-1.  
 GN ETS-1.  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88300873.  
 RA LEPINE D., DUTERQUE-COQUILLAUD M., LI R.-P., HENRY C., FLOURENS A.,  
 RA DEBUIRE B., STEHELIN D.;  
 RT "Alternative splicing within the chicken c-ets-1 locus: implications  
 RT for transduction within the E26 retrovirus of the c-ets  
 RT proto-oncogene.";  
 RL J. VIROL. 62:3233-3241(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88206088.  
 RA WATSON D.K., MCWILLIAMS M.J., PAPAS T.S.;  
 RT "Molecular organization of the chicken ets locus.";  
 RL VIROLOGY 164:99-105(1988).  
 CC -1- FUNCTION: THIS PROTEIN IS THE NORMAL CELLULAR PRODUCT OF CHICKEN  
 CC ETS. IN THE E26 VIRUS, ETS IS RESPONSIBLE FOR ERYTHROBLAST  
 CC TRANSFORMATION.  
 CC -1- TISSUE SPECIFICITY: SPLEEN.  
 CC -1- ALTERNATIVE PRODUCTS: ALTERNATIVE SPLICING MECHANISM WITHIN THE  
 CC C-ETS-1 GENE GENERATES TWO PROTEINS HAVING DISTINCT N-TERMINAL  
 CC EXTREMITIES: P54 AND P68.  
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M21823; G211466; .  
 DR EMBL; M20515; G211464; ALT\_INIT.  
 DR EMBL; M20507; G211464; JOINED.  
 DR EMBL; M20508; G211464; JOINED.  
 DR EMBL; M20509; G211464; JOINED.  
 DR EMBL; M20510; G211464; JOINED.  
 DR EMBL; M20511; G211464; JOINED.  
 DR EMBL; M20512; G211464; JOINED.  
 DR EMBL; M20513; G211464; JOINED.  
 DR EMBL; M20514; G211464; JOINED.  
 DR PIR; A28875; TVCHET.  
 DR PROSITE; PS00345; ETS\_DOMAIN\_1; 1.  
 DR PROSITE; PS00346; ETS\_DOMAIN\_2; 1.  
 DR PROSITE; PS00346; ETS\_DOMAIN\_2; 1.  
 DR PFAM; PF00178; Ets; 1.  
 DR HSSP; P27577; 1ETD.  
 DR TRANSFAC; T00115; .  
 KW PROTO-ONCOGENE; DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.  
 FT DNA\_BIND 379 459  
 FT CONFLICT 220 220 S->C (IN REF. 2).  
 FT CONFLICT 222 222 P->R (IN REF. 2).  
 SQ SEQUENCE 485 AA; 55021 MW; C58D8897 CRC32;

Query Match 15.28; Score 272; DB 1; Length 485;  
 Best Local Similarity 24.28; Pred. No. 6,4e-14;  
 Matches 112; Conservative 42; Mismatches 109; Indels 200; Gaps 18;

QY 65 MLIPEDSSMAKAP- -GASSREPEEP- -EQCPYIDSA- -101  
 DB 2 MSYMDTIGSTGPYPLARGVWGAGSSCCEDPMMCRQSACCPRCPPMDNAIOE 61  
 QY 101 -PAG- -SIDLYPGGLTLEHSLLEVQSMVYGEVAKDIETACKLINTADPMDSPSNV 155  
 DB 62 VPGLEHYSTDMECAVPLTTPSKKMSQALKATFSGRPKDQRIKIPKQWTEHY 121  
 QY 156 QKWLNTHEOYRLPMPKAFQ- -LAGKELCAMSEDFRORSP-LGGVLAHLIDK- -211  
 DB 122 RDVMAVAVNEFSL- -KGVDQFKCMGALCALGKCEFLAPDVGILHEHLEILOKE 179  
 QY 211 - -SAANKER-TSPGAIHY- -CASTSE- -ESW- -236  
 DB 180 EAKPYANGVNAAYPESRTSDYFISGIEHAQVCPSESEPSFTSOTLHPISSEE 239  
 QY 236 - -TDS- -EV- -DSSCSQ- -248  
 DB 240 LLSKRYENDYPSVILRDYQVDTSDTDFTYIKQEVYTDNMCNGASGKLGQDSESI 299  
 QY 248 - -248  
 DB 300 ESYDSCDRLTQSSQSSFSQSLQRPYSDFSEDPALPNHKKPGTFKDYVRADMN 359  
 QY 248 - -PIHLMQFLKELKPHSYGFIWLNKKEGIFRIEDSAOVAR 289  
 DB 360 KDKPYIPAAALAGYGSFPQIQLWQFLE-LTLDKSCGSFISWTG-DWEKRLSDPEVAR 417  
 QY 290 LWGIRKNRPANNDKLSRSIROYKKGIIKRPDISQRLVYO 332  
 DB 418 RGRKRNKPKMYEKLISGLRYDDKNIIHK-TAGKRYVAFV 459

RESULT 10  
 FLII\_MOUSE STANDARD; PRT; 452 AA.  
 AC P26323;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)



RETRORVIRAL INTEGRATION SITE PROTEIN FLI-1.  
 FLI1 OR FLI-1.  
 OS MUS MUSCULUS (MOUSE).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-SPLEEN;  
 RA MEDLINE: 91257578.  
 RA BEN-DAVID Y., GIDDENS E.B., LETWIN K., BERNSTEIN A.;  
 RT "Erythroleukemia induction by Friend murine leukemia virus:  
 RT Insertional activation of a new member of the ets gene family, flt-1,  
 RT closely linked to c-ret-1".  
 RL GENES DEV. 5:908-910(1991).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE: 93275657.  
 RA ZHANG L., LEMARCHAND V., ROMEO P.-H., BEN-DAVID Y., GREER P.,  
 RA BERNSTEIN A.;  
 RT "The flt-1 proto-oncogene, involved in erythroleukemia and Ewing's  
 RT sarcoma, encodes a transcriptional activator with DNA-binding  
 RT specificities distinct from other Ets family members.";  
 RL ONCOGENE 8:1621-1630(1993).  
 CC -1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES  
 CC THE DNA SEQUENCE 5'C(CA)GGAGT-3'.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH HEMATOPOIETIC AND  
 CC NONHEMATOPOIETIC TISSUES.  
 CC -1- INDUCTION: SPECIFICALLY UP-REGULATED IN LEUKEMIC CLONES WITH  
 CC F-MULV INSERTIONS UP-STREAM OF THE FLI-1 LOCUS.  
 CC -1- DISEASE: INVOLVED IN ERYTHROLEUKEMIA INDUCTION BY FRIEND MURINE  
 CC LEUKEMIA VIRUS (F-MULV).  
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.  
 CC .....  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as the content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC .....  
 DR EMBL: X59421; G50975; .  
 DR PIR: S17403; S17403.  
 DR MGI: 95554; FLI1.  
 DR PROSITE: PS00345; ETS\_DOMAIN\_1; 1.  
 DR PROSITE: PS00346; ETS\_DOMAIN\_2; 1.  
 DR PROSITE: PS00061; ETS\_DOMAIN\_3; 1.  
 DR PFAM: PF00178; Ets; 1.  
 DR HSP: Q01543; 1FLI.  
 DR TRANSFAC: T01408; .  
 KM TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;  
 KM PROTO-ONCOGENE.  
 FT DNA\_BIND 281 361 ETS-DOMAIN.  
 FT SEQUENCE 452 AA; 51002 MW; CECACDDA CRC32;  
 Query Match 15 28; Score 272; DB 1; Length 452;  
 Best Local Similarity 25.6%; Fred. No. 5.8e-14;  
 Matches 102; Conservative 46; Mismatches 142; Indels 108; Gaps 16;  
 Oy 2 GSASGLSSVSPSHLLPPDVTGRTGLEKAAAGVGLERDRSPSPATPEOG-----LSA 57  
 Db 3 GIKKALSVSDOSL-----FDSAYAAAHLPADMTAS--GSPDGGORHKINP 50  
 Oy 58 FYLSYFDMLYPEDSSMAKAPGASSRE-----PPEBPQCPIVDISQAPAGSLDLYPGC 111  
 Db 51 L-----PQGEIMNPVAVNKKREYDHNGSRSPVDSYKCNKLVGGGANPMN 101  
 Oy 112 LT--LEHSLEVOSSVWVGEVLKDIETACKLNTADPMDSPSPVOKMLTMEHQYRLP 169  
 Db 102 YNSYDEKNGPPPMNTNE-----RVIIVPADPTLTQEHVRQMLEVAIKYEGLM 152

170 PMGRA-FOELAGKELCAMESEOF--RORSPGLGVDYLAHL----- 207  
 Db 153 EIDTSFONMGKELCKKNKEDFLRATSAVNTYEVLLSHLSLRESSLLAYTTSHITDOSS 212  
 Oy 207 -----DIMSAAV-----MKERTSPGAHYCASTSEE 233  
 Db 213 RLNVKEDPSYDSVARGANNMNSGLNKPPLLGSSQTKTKTEQRPQDPYOILGPTS-- 271  
 Oy 234 SWTSEVDSSCGGPPIHMLQFLKELLPHSYGRFIRNLNKEGIFKIEDSAQVRLMGI 293  
 Db 271 ----SRLNPPSGQ-IQLMOPLE-LLSDSANASCITW-EGTNGEFKMTDDEVARRWGE 323  
 Oy 294 RKNRPANYDLSRSIRORYKKGIIRKPDISQRLVYOF 331  
 Db 324 RRSKPNMYDKLSRLRLRYIDKNIMTKVH-GKRYAYKF 360  
 RESULT 11  
 FLI1\_HUMAN  
 ID FLI1\_HUMAN STANDARD; PRT: 452 AA.  
 AC 001543;  
 DT 01-JUL-1993 (REL. 26, CREATED)  
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE FLI-1 ONCOGENE (ERGB TRANSCRIPTION FACTOR).  
 GN FLI1.  
 OS HOMO SAPIENS (HUMAN).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-BONE MARROW;  
 RX MEDLINE: 92396239.  
 RA DELATRE O., ZUCMAN J., PLOUGASTEL B., DESMAZE C., MELOT T.,  
 RA PETER M., KOVAR H., JOGBERT I., DE JONG P., ROULEAU G.;  
 RT "Gene fusion with an Ets DNA-binding domain caused by chromosome  
 RT translocation in human tumours.";  
 RL NATURE 359:162-165(1992).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 93075640.  
 RA WATSON D.K., SMYTH F.E., THOMPSON D.M., CHENG J.O., TESTA J.R.,  
 RA PAPAS T.S., SEH A.;  
 RT "The ERGB/Flt-1 gene: Isolation and characterization of a new member  
 RT of the family of human Ets transcription factors.";  
 RL CELL GROWTH DIFFER. 3:705-713(1992).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 93007976.  
 RA PRASAD D.D., RAO V.N., REDDY E.S.;  
 RT "Structure and expression of human flt-1 gene.";  
 RL CANCER RES. 52:5833-5837(1992).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-BLOOD;  
 RC MEDLINE: 93167799.  
 RA HROMAS R., MAY W., DENNY C., RASKIND W., MOORE J., MAKI R.A.,  
 RA BECK E., KLEMSZ M.J.;  
 RT "Human flt-1 localizes to chromosome 11Q24 and has an aberrant  
 RT transcript in leuopeleloma.";  
 RL BIOCHIM. BIOPHYS. ACTA 1172:155-158(1993).  
 RN [5]  
 RN STRUCTURE BY NMR OF 276-373.  
 RP MEDLINE: 95292091.  
 RA LIANG H., MAO X., OLEJNICZAK E.T., NETTESHEIM D.G., YU L.,  
 RA MEDDOWS R.P., THOMPSON C.B., FESIJK S.W.;  
 RT "Solution structure of the ets domain of flt-1 when bound to DNA.";  
 RL NAT. STRUCT. BIOL. 1:871-875(1994).  
 CC -1- FUNCTION: SEQUENCE SPECIFIC TRANSCRIPTIONAL ACTIVATOR. RECOGNIZES  
 CC THE DNA SEQUENCE 5'C(CA)GGAGT-3'.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- DISEASE: A FORM OF EWING'S SARCOMA IS CHARACTERIZED BY A  
 CC CHROMOSOMAL TRANSLOCATION T(11;22)(Q24;Q12) WHICH INVOLVES FLI1

```

CC      AND EMS.
CC      -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X67001; G32530;
DR      EMBL: M98833; G182189;
DR      EMBL: S45205; G257354;
DR      EMBL: M93255; G182660;
DR      PDB: 1FLI; 15-SEP-95.
DR      MIM: 193067;
DR      PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR      PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR      PROSITE: PS0061; ETS_DOMAIN_3; 1.
DR      PFAM: PF00178; Ets; 1.
DR      TRANSFAC: T02066;
DR      TRANSCRIPTION REGULATION: ACTIVATOR; DNA-BINDING; NUCLEAR PROTEIN;
KW      PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION; 3D-STRUCTURE.
FT      DNA_BIND 281
FT      CONFLICT 69
FT      CONFLICT 77
FT      CONFLICT 130
FT      CONFLICT 133
FT      CONFLICT 323
FT      CONFLICT 391
FT      CONFLICT 426
SQ      SEQUENCE 452 AA; 50962 MW; 192476f7 CRC32;

Query Match
Best Local Similarity 15.1%; Score 270; DB 1; Length 452;
Matches 103; Conservative 50; Mismatches 137; Indels 108; Gaps 17;

QY      2 GSASGLSSVSPSHLLPPDYVTSRTGLEKAAAGAVGLRRDMSPPAPTPBOG-----LSA 57
DB      3 GIKKALSVSDOSL-----FDSAGAAAHLPKADMTAS--GSPDYGGPHKINP 50
QY      58 FYLSYFDMLYEDSSMAKAFGASRRE-----PPEEPCPYIDSOAPAGSLDLYPGG 111
DB      51 L-----PPOQEMINQPVAVYKREYDMMNSRSPVDCSY-----SKCKLVGGG 95
QY      112 LT-----LEHSLDEVQSNVGEVLKDIETACKLLNITADPMDSPPSNVQKMLMTE 163
DB      96 ESNPNVNSYNDENKGPPEPPMTTNE-----RRIYVADPTLWTQEHVRQWLEMAI 146
QY      164 HOYRLPPMKKA-FOELAGKEICAMSEBOFRORSPL-GGDVYLAHL----- 207
DB      147 KEYSLEMDITFFQNMCKELCKMKNEDEFLRATLYINTEVLSHLSYLRSSLLAYNTTS 206
QY      207 -----DIKSAAMKERTSP-----GAHYCASTSE----- 234
DB      207 HTDSSRLSVKEDPSYDVRGAWGNMNSGLNKSPLLGAGQITSKNTEBRPQDPYQIL 266
QY      234 SWTSEVDSSCGPPIHLMQFLKELLKPHSYGRFIMLWLNKKGIFKIEDSAOVARLMGI 293
DB      267 GPTSSRLANPSSGO-IQLMQFLLE-LLSDSANASCITW-EGTNGEFTMDPDEVARRMGE 323
QY      294 RKNRPANMYDLKRSIRQYKGIIRKPDISOQLVYOF 331
DB      324 RKSCKPNMYOKLSRALRYIYDKNIMTYVH-GKRIATKF 360

```

```

DT      15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE      15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE      ETS-RELATED PROTEIN TEL (ETS TRANSLOCATION VARIANT 6).
CN      ETV6 OR TEL.
OS      MUS MUSCULUS (MOUSE).
OC      EUDAROTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC      RODENTIA; SCIROGNATHI; MORIDAE; MORINAE; MUS.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RA      BERNARD O.;
RL      SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC      -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Y07915; E274398;
DR      MGD: MGI:109336; ETV6.
DR      PROSITE: PS00345; ETS_DOMAIN_1; FALSE_NEG.
DR      PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR      PROSITE: PS0061; ETS_DOMAIN_3; 1.
DR      PFAM: PF00178; Ets; 1.
DR      HSSP: Q01543; 1FLI.
KW      NUCLEAR PROTEIN; DNA-BINDING.
FT      DNA_BIND 335
FT      DNA_BIND 416
SQ      SEQUENCE 485 AA; 56405 MW; C1A75D87 CRC32;

```

```

Query Match
Best Local Similarity 14.9%; Score 265.5; DB 1; Length 485;
Matches 98; Conservative 47; Mismatches 80; Indels 223; Gaps 15;

QY      75 AKAPASSRRE-----PPEEPCPYIDSOAPAGSLDLY-----VGGLLLEHSLDEVQSNV 126
DB      2 SETPAOSSIKQERISTYTPESP-----VASHRSTPLVHVHPALRREEDSIH----- 51
QY      127 VGEVLNDIETACKLLNITADPMDSPPSNVQKMLMTEHOYRLPPMKAFQELAGKEICAM 186
DB      51 -----LPTHLRL-----QPIYMSRDVDAQMLKAENEFSLRPLESKFEMNGKALLL 98
QY      187 SEQAFRQRPPLGGDVH---AHLDIKSAAMKERTS-----PG----- 223
DB      99 TKEDFRKSPHSGDVLYELLOHI-----LQKRSMLFSPFPFQDSIHTKPEVLLH 150
QY      223 -----AIHYCASTSE----- 233
DB      151 QNHDEDCVQRTPRTPASVHNHPPTIELLRPSPTITNHRPDPDQORPORSPIDNM 210
QY      233 -----ESMTDS----- 239
DB      211 SRRLSVPEKAGPRLQDENNHQETYPILSVPENNHCLPSSPQGESTRVYQIMSPITMH 270
QY      239 -----EYDSSCS-----GPIH----- 251
DB      271 LILNRHSHSVDFKQSRSEDEGANNREGPINLSHREDLAVLNHIMVMSPEERAMPGR 330
QY      251 -----LMQFLKELLKPHSYGRFIMLWLNKKGIFKIEDSAOVARLMGIKKNPAMMYDK 304
DB      331 IADCRLLMDVYQ--LLSDSREYENIRNEDKESKIFRIYDNGLARLMGNKKNRTNMYEK 389
QY      305 LSRISRIQYKGIIRKPDISOQLVYOFV 332
DB      390 MSRALRHYKLIINIRK-EPGQRLFRFM 416

```

RESULT 13



```

RT Protein.
RL Virology 167:1-7(1988).
CC - FUNCTION: THIS PROTEIN IS THE NORMAL CELLULAR PRODUCT OF CHICKEN
CC ETS IN THE E26 VIRUS, ETS IS RESPONSIBLE FOR ERYTHROBLAST
CC TRANSFORMATION.
CC - TISSUE SPECIFICITY: THYMUS.
CC - ALTERNATIVE PRODUCTS: ALTERNATIVE SPLICING MECHANISM WITHIN THE
CC C-ETS-1 GENE GENERATES TWO PROTEINS HAVING DISTINCT N-TERMINAL
CC EXTREMITY: P54 AND P68.
CC - SIMILARITY: BELONGS TO THE ETS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M22462; G211753;
DR EMBL: X13026; G63180;
DR EMBL: X13027; G63383;
DR PIR: A31285; TVCHTE.
DR PIR: S07625; S07625.
DR PIR: S29132; S29132.
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR PFAM: PF00178; Ets; 1.
DR HSSP: P27577; IETD.
DR TRANSFAC: T00114;
DR PROTO-ONCOGENE: DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING;
KW PHOSPHORYLATION.
FT DOMAIN 54 135 POINTED.
FT DNA_BIND 335 415 ETS-DOMAIN.
FT MOD_RES 38 38 PHOSPHORYLATION (BY MAPK) (BY
FT SIMILARITY).
FT CONFLICT 117 117 R -> A (IN REF. 3).
FT SEQUENCE 441 AA; 50326 MW; 734E39B3 CRC32;
SO

```

```

Query Match 14.18; Score 252.5; DB 1; Length 441;
Best Local Similarity 24.08; Pred. No. 1.8e-12;
Matches 96; Conservative 38; Mismatches 89; Indels 177; Gaps 15;

```

```

QY 105 LDVLP-----GGLTLEHSLEOVSMYGEVLKDTACKLLNTADPMDWSPSNVQKW 158
DB 21 IDLPSPDMECADVPPLTPPSKEMMSQALKATFSGFAKEQRLGIPDPQOWTEHYRDM 80
QY 159 LMTHEGYRLPMPGKAFQ--LAGKELCAMSEOFQRSP-LGGDVLAHLIDTWK----- 211
DB 81 VMAMVNEFSL--KGVDFQKCMNGALCALGKCEFLERDVEVDILWEHLLETLQREAK 138
QY 211 -----SAAMMER-TSPGAHRY-----CASTSE-----ESW----- 236
DB 139 PYPMGVNAAYPESRYSDYFISYGIHACVPPSESEPSFTESQYLRHPISSELLS 198
QY 236 -----TDS-----EV---DSSCSGQ----- 248
DB 199 LKTEYDPSVILDDPVOTDLSQTDYFTIKQEVYTPDMCMGRASRGKLGQDSFESIESY 258
QY 248 ----- 248
DB 259 DSCDRLTQSWSSQSFQSLQRPVSDYDFSEDYPALPMPHKKRGCTKYVYRDRADNMKK 318
QY 248 -----PIHMQFLKELLKPKHSGYGRIRMLNNEKGIJFKIEDSAQVARKLWG 292
DB 319 PVIPAAALAGYTGSGPIQIMQFL--LLTDKSCOSFISWTG--DGWFEKLDSDDEVARWAG 376
QY 293 IRKNRPAMVYDKLSRIROYKKGIIRKPDISORLYQYQV 332
DB 377 KRKNKPKMYEKLSRGLRYTYKNIHK--TAGKRYRYREV 415

```

```

RESULT 15
ET2B_XENLA
ID ET2B_XENLA STANDARD; PRT: 472 AA.
OC 091712;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE C-ETS-2B PROTEIN.
GN ETS2B OR ETS-2B.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOATRACHIA; PIPOIDEA; LIPIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-OVARY;
RC MEDLINE: 90356411.
RA WOLFE C.M., STEIGLER P., BALTZINGER M., MEYER D., GHYSDAEL J.,
RA STEHLEIN D., BEFORT N., REMY P.;
RT "Isolation of two different c-ets-2 proto-oncogenes in Xenopus
RT laevis."
RL NUCLEIC ACIDS RES. 18:4603-4604(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-OVARY;
RC MEDLINE: 92088972.
RA WOLFE C.M., STEIGLER P., BALTZINGER M., MEYER D., GHYSDAEL J.,
RA STEHLEIN D., BEFORT N., REMY P.;
RT "Cloning, sequencing, and expression of two Xenopus laevis c-ets-2
RT proto-oncogenes."
RL CELL GROWTH DIFFER. 2:447-456(1991).
CC - SUBCELLULAR LOCATION: NUCLEAR.
CC - SIMILARITY: BELONGS TO THE ETS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X52635; G64621;
DR PROSITE: PS00345; ETS_DOMAIN_1; 1.
DR PROSITE: PS00346; ETS_DOMAIN_2; 1.
DR PROSITE: PS50061; ETS_DOMAIN_3; 1.
DR PFAM: PF00178; Ets; 1.
DR HSSP: P14921; 25TW.
KW DNA-BINDING; NUCLEAR PROTEIN.
KW DNA_BIND 366 446 ETS-DOMAIN.
FT SEQUENCE 472 AA; 53928 MW; 196506B4 CRC32;
SO

```

```

Query Match 14.08; Score 250; DB 1; Length 472;
Best Local Similarity 22.98; Pred. No. 3.1e-12;
Matches 92; Conservative 35; Mismatches 91; Indels 184; Gaps 15;

```

```

QY 108 VPGGLTLEHS-----LEOVSMYGEVLKDTACKLLNTADPMDWSPSN 154
DB 52 VPTGLDSTPHSSSCCEPLPLTPCSKAVMSQALKDTENGFAKERCR-LGIPGNPWLMDENN 110
QY 155 VOKMLTMEGYRLPMPGKAFQ--LAGKELCAMSEOFQRSP-LGGDVLAHLIDTWK----- 208
DB 111 VLOWLWMAKKEFSLENNV--FQKFLNMGHCLSGIKERFLALAPDPVGDILWEHLEEMAK 168
QY 208 -----IKMSAMMERKTRSP----- 222
DB 169 EYQKAQEPYITDSNRDSLQNMNADSLNFTADPLOGAOVHNYPRNGMYNDKCSVPTGO 228
QY 222 -----GAHYCASTSE----- 233
DB 229 TLNPKQEFQOYDPSSCLSKSRVAVYPPASODFARSHMNVLLNSLNGKLDYDSDGSGTES 288

```

Search completed: November 20, 1999, 15:17:38  
Job time: 57 sec

**THIS PAGE BLANK (USPTO)**